DOCUMENT RESUME

ED 291 759 TM 011 044

AUTHOR Hoffman, Lee McGraw
TITLE Spoonfeeding, Handfeeding, Forcefeeding, and Feedback: Identifying Manipulable Variables in Evaluation Utilization. Definitions and Models in Evaluation Use.
PUB DATE Oct 87
NOTE 18p.; Paper presented at the Annual Conference of the American Evaluation Association (Boston, MA, October 16, 1987).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Evaluation Methods; *Evaluation Utilization; Models; Research Utilization
IDENTIFIERS *Descriptive Method; Manipulable Variables

ABSTRACT

This paper attempts to define evaluation use and manipulable variables related to it. Approaches to defining evaluation use have included typologies inventorying kinds of evaluation use, threshold criteria defining accepted levels of evidence, construct or operational definitions, and dependence on the perspective of the decision maker. None of these approaches is found to be satisfactory. Within models of evaluation use, the concept is conceived as a dependent variable, and much of the literature constitutes an attempt to identify the factors that act as independent variables in their relation to use. Descriptive models are often structured around components of an evaluation and its setting. Explanatory models, of which there do not seem to be many in the literature of evaluation use, include those developed by A. J. Meltsner (1976), R. D. Brown et al (1984), and G. E. Hall (1981). An effective model of manipulation should have the following characteristics: (1) open-mindedness in defining use as well as non-use and misuse; (2) conceptualization of evaluation use as a process; (3) inclusion of relevant components and their interrelationships; (4) treatment of manipulable and non-manipulable variables; (5) central focus on the interaction of factors; and (6) reduction of variables as much as possible. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED291759

SPOONFEEDING, HANDFEEDING, FORCEFEEDING, AND FEEDBACK: IDENTIFYING MANIPULABLE VARIABLES IN EVALUATION UTILIZATION

Definitions and Models in Evaluation Use

Lee McGraw Hoffman
Department of Adult and Community College Education
North Carolina State University
Raleigh, North Carolina 27695-7801

Paper presented at the annual conference of the American Evaluation Association, Boston, Massachusetts, October 16, 1987.

TM011 044

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lee McGraw Hoffman

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

DEFINITIONS AND MODELS IN EVALUATION USE

Lee McGraw Hoffman
Paper presented at the meeting of the American Evaluation Association, Boston, October 1988

Introduction

Two points need to be made about this paper at the outset. The first is that it grew circuitously, directed by questions and confusions that arose from attempts to write something quite different. The structure follows the meandering path set by questions, tentative answers, and new questions that these opened up. The second point is that this is intended as a prompt for discussion. None of the questions is resolved, the review of the literature is nowhere near exhaustive, and the ideas here are presented for challenge.

Problem

This paper began as an exercise in testing a model of factors associated with evaluation use (Cousins and Leithwood, 1986) by examining the model's adequacy in explaining two evaluations that had many contextual similarities but that differed considerably in the extent to which they were used. It was possible to describe these evaluations in terms of Cousins' and Leithwood's clusters of factors. There were also differences in the situa tions, actions, etc. grouped under these factors that were plau sible reasons for one study's being shelved while the other ended with the program's field staff spontaneously adopting on going evaluation procedures. But the exercise was less than

satisfactory for several reasons:

1) There was no conceptual linkage between the kinds of use in the two studies and factors affecting use; in other words, it was not possible to hypothesize what conditions would lead to what kind of use;
2) While it was possible to distinguish between high- and low-use evaluation cases, it was not possible to reduce the explanation to a form more concise than the list of factors with which the cases were described;
3) The exercise yielded few insights about how the study or the context in which it took place could have been manipulated to affect use.

These problems should not be attributed to the Cousins and Leithwood model. Those authors note the need for further research to determine whether different kinds of evaluation use are associated with differences within the factors they report, and to test the higher-order factors they suggest as organizers of the dozen variables they identify empirically.

However, this sense of unease led to an attempt, begun as a mechanical exercise, to outline the concept (evaluation use) and the relevant principles (factors) in a model that explained and predicted varying kinds and degrees of use (manipulable variables). That exercise was unsuccessful and led, in turn, to several questions on how we think about evaluation use:

1) How may, or do, we define evaluation use?
2) What kinds of models of evaluation use do we have?
3) What do answers to the preceding questions tell us about manipulating evaluation use?

Defining Evaluation Use

Defining evaluation use is more than an academic issue. Considerations of nonuse have included concerns about methodological weaknesses in evaluations, instances of deliberate misuse, and, with increasing frequency, the suspicion that evaluators and those who use their products have different ways of looking at the world (Burry, 1985; Patton, 1980; Rutman, 1980). Evaluators live in a world of simple, linear, immediate cause-and-effect. Policy makers live in a world of complex interactions among political, organizational, and professional constraints (Weiner, Rubin and Sachse, 1977). Change is a process of disjointed incrementalism, to which evaluation results contribute in gradual, cumulative, indirect ways (Ruskus and Alkin, 1984). Since policy makers continue to pay evaluators for their work, it seems likely that some reported nonuse is an artifact of differences in how the two groups define "use." Within evaluation literature, apparent nonuse may also be due to evaluators insisting upon a restrictive definition of the phenomenon (Wise, 1977).

There has been considerable discussion about defining use. Ciarlo (1981) identifies a range of definitions that fall into two major camps, the traditional instrumental (overt use) definition and the more recent identification of use as conceptualization (changes in thinking). He notes also that most investigators study only one type of use or fail to differentiate between factors expected to produce one or the other, and that there is often an assumption that study results will hold for some gene-

ral, undifferentiated use. Weiss echoes this when she stresses that the field will face noncomparable studies until the question of defining use has been resolved (1981). The question of definition may remain unresolved, however, because we have not yet determined a universally accepted definition of evaluation (Patton, 1985). It may be an unresolvable issue because as long as evaluation serves a variety of purposes we can expect diversity in studying how its results are used (Merwin and Wiener, 1985).

Approaches to Defining Use

If there is no accepted single definition of evaluation use, there are several identifiable approaches to the task. These are typologies, descriptions of threshold criteria, constructs and operational definitions, and decision oriented approaches.

Typologies inventory kinds of evaluation use, sometimes as points on a continuum, but more often as straightforward lists. Weiss (1981) argues for a continuum ranging from conceptual to instrumental use, with many "middle ground" cases in which the evaluation is part of, but does not determine, the decision. Leviton and Hughes (1981) see three categories of use emerging from other literature. These include instrumental use of an evaluation in specific ways in a decision; conceptual use that influences thinking without specific, documentable results; and persuasive (sometimes 'symbolic') use that calls upon evaluation results to bolster personal influence in defending a position or convincing others of a plan of action. King and Thompson (1981) identify the same categories of use in their literature review. Several years later, Cousins and Leithwood (1986) analyze empiri-

cal research on evaluation use and find it possible to reduce the definitions to three categories: influencing specific decisions, educating users, and the psychological processing of evaluation information.

Some writers approach definition by specifying threshold criteria-- use exists if acceptable levels of evidence can be found. Alkin, Daillak and White (1979) maintain that there are four necessary conditions for evaluation use: information must have been communicated, a user must be evident, information must be labeled as an input, and the information must be used in some way. There are two "bottom line" conditions specified by Leviton and Hughes: evidence of serious discussion that translates the findings into their implications, and evidence that people would have thought or acted differently in the absence of the evaluation results (1981).

There are definitions that address evaluation use as a construct or operational definition. These generate definitions of, or boundaries for, evaluation use from underlying concepts and dimensions. In some cases the emphasis is on limitation. Hall (1981), for example, suggests that evaluators err in confusing evaluation use with its consequences, much as Leviton and Hughes (1981) warn that evaluation use should not be mistaken with impact (the outcome of use) or utility (relevance of the findings). King and Pechman (1984), however, present evaluation use as a subset of impact because they define use as as an intentional act directed toward a specified end. These authors also diagram the interrelationships of various kinds of use and

describe the definitionally necessary conditions for several of these kinds. Guba and Lincoln (1981) develop a construct of four basic evaluation uses by combining the two dimensions of merit/worth and formative/summative. Patton (1985) similarly separates the two dimensions of use/nonuse and misuse/nonmisuse to discuss the various combinations that can occur.

The final major kind of definition is that which is implied in writings that discuss evaluation use from the perspective of the decision makers. Although this literature usually does not have definitions in the sense of, "evaluation use is...," working definitions can be inferred from the discussion of the evaluation's role in the decision making process. Patton describes evaluation utilization as a "diffuse and gradual process of reducing decision maker uncertainty within an existing social context," in a manner that feeds into "the slow, evolutionary process of program development" (1980, p.273). This reflects Kleinfeld and McDiarmid's arguments that evaluations should lessen the distance between scientific objectivity and decision makers (1986) and Shapiro's advice that evaluations should be tailored around decison makers and integrated with the existing decision process (1985). Ruskus and Alkin (1984) also extend their definition of use to include the broad, supporting role evaluation findings sometimes play in the decision making process, their incremental influence. In these writings evaluation is seen as an ongoing component of the decision making process and the evaluation itself is not the sole, nor even the central, focus. In fact, it is tempting to categorize these as "ecological" definitions because their purpose is to describe the evalua-

tion as a functioning part of its natural environment.

Reactions to Definitions

None of the four kinds of categorization discussed is completely satisfactory. Typologies have the advantage of being generally grounded in observation, but often what is observed is evaluation's gradual weaning from the classic experimental model. Furthermore, it is difficult to tell at this point whether the same types recur because they are actually there or because they have influenced the research on evaluation use.

Establishing threshold criteria implies that use has been defined: saying how much of something is needed to register on our scale means that the "something" exists in identifiable form. By themselves, these criteria are sterile. Constructs are intuitively most attractive but the breakout cited from Guba and Lincoln really addresses purpose (value, goal) rather than use (actual employment). King and Pechman develop a taxonomy of use that is potentially helpful; the interrelationships of the different kinds of use they list and the question of whether user intent is really required in a definition need to be discussed further. The other constructs reported here are more in the nature of caveats against confusing issues than they are first principles that can serve to generate a theory.

If the purpose of this discussion is to describe evaluation use in such a way that it is amenable to manipulation, then writings concerned with its role in the decision making process appear most helpful. This approach focuses on the evaluation's environment and on the factors that are usually considered out-

side the evaluator's span of control. It is an open perspective as opposed to the "closed system" approach of treating an evaluation as a self-contained event, born in the problem statement and ending when its recommendations are apotheosized in program change. Looking at the decision maker also encourages us to see evaluation use as a process that occurs between the evaluation and the user, and emphasizes the situational context. This discourages reification of evaluation use, the error of treating use as if it had some existence independent of the user and the situation.

The categories identified so far can be seen as byproducts of the debate that has been ongoing as evaluation matures (or declines, depending on your position) from its original ties to the traditional experimental research model. At this point there does not seem to be one acceptable, all-purpose definition of evaluation use. Such a definition should partake of each category. It should describe evaluation use, specify the criteria for inclusion of a case in the definition, and present a rounded concept strong enough to support a model of evaluation use. The following discussion argues that the absence of a definition meeting these conditions has some implications for model development.

Models of Evaluation Use

Evaluation use is typically conceived as a dependent variable, and much evaluation use literature is an attempt to identify the factors that act as independent variables in their relation to use. Once these variables have been determined,

however, we need a construct that expresses this relationship. It is one thing, for example, to say that mosquito populations are related to temperature, rainfall, and season. It is another to explain this relationship through the mosquito's breeding cycle.

What is desired is a model of evaluation use that is exhaustive, parsimonious, and has explanatory power. Again, cursory review of the literature on evaluation use reveals two types of models. The first, and considerably more common, describe factors related to evaluation use. The second category attempts to explain the relationships.

Descriptive Models

Descriptive models often are structured around the components of an evaluation and its setting. Holley (1979) writes that utilization factors must consider the charactersitics of the evaluand, the users, the organization, the evaluator, the findings, and the dissemination resources. These are simplified into broad groupings of human, context, and evaluation factors (Ruskus and Alkin, 1984) or personal, organizational, and procedural or methodological variables (Burry, 1985; Haenn, 1980). Not only the factors, but their interrelationships are important, as in the examination of evaluator, audience, and report characteristics cited in Braskamp and Brown (1980).

Often these components of an evaluation serve to organize lists of factors related to use. Descriptive lists are seen in research such as that of Cousins and Leithwood (1986), which

proposes twelve factors (e.g., evaluation quality, information needs) and two higher order factors of 1) evaluation implementation and 2) the decision or policy setting. Leviton and Hughes (1981) propose five clusters of variables, such as relevance, communication, and user involvement and adequacy. The same factors appear in prescriptive lists. Carver (1985) advises that evaluations will be used when enough resources are put into them (when they are not isolated and underfunded), the time for impact is right, and decision makers have commissioned the study. Kleinfeld and McDiarmid (1986) suggest practical strategies for the varying phases of an evaluation that reflect a concern with relevance, user involvement, evaluator credibility, and considered dissemination.

Many of the descriptive (or prescriptive) models include the same factors. Some approach explanatory status, as in an article by Siegel and Tuckel (1985) that presents factors as they mediate between the conduct of an evaluation and the use of its evaluation findings. It should also be noted that the separation of descriptive models into the subcategories discussed here is arbitrary. Writers typically cross aspects of evaluation components (actors, phases, and the like) with lists of relevant factors to generate prescriptions about what should be done to encourage use.

Explanatory Models

There does not seem to be a great concern with explanation in the literature of evaluation use. From a policy perspective, Meltsner (1976) suggests that use can be explained in large

part by the evaluator's political and analytical skills, and develops a typology of four evaluator profiles. This is an explanatory model of sorts.

Brown, Newman and Rivers (1984) present two decision making models and derive from these three factors that influence use: perceptions of risk, time limitations, and whether there is a solution available. High and low levels of these factors are then combined to produce four possible situations and the probable decision maker response under each situation is discussed. This is explanatory and also has a prescriptive aspect as it portrays the evaluator's concerns in each situation.

Hall (1981) argues that evaluation use can be seen as an innovation and can be described and explained in his levels of use framework. This idea is expanded by Pechman and King (1986), who detail the adoption of evaluation as an innovation within local education agencies over a continuum progressing from nonuse to renewal.

Viewing evaluation use as an example of innovation has the advantage of providing a parallel with a clear and tested model of adoption. It is a little difficult, however, to tell if evaluation use is a process synonymous with innovation or with innovation adoption, or if it is an innovation to be adopted. Each interpretation has widely different consequences in establishing a model for evaluation use. The latter, for instance, results not so much in a model as it does in a recounting of the steps to be undertaken.

The Brown, Newman and Rivers model may be an adequate one.

It should be examined to see how well it accounts for or subsumes the factors identified in descriptive models. Can, for example, time, risk, and availability of a potential solution explain such things as communication clarity, the existence of conflicting findings, and the information needs of the audience? Upon immediate reading this model appears much more likely to explain the decision setting cluster identified by Cousins and Leithwood than it does their cluster of evaluation implementation factors.

Manipulating Evaluation Use

Much of what has been reviewed here-- and a great deal more evaluation literature that has not been cited-- provides solid, practical advice on how to enhance the use of evaluation products and processes. It really does help to collaborate with clients in designing studies, to give preliminary informal briefings so that stakeholders can adjust to unexpected or unwelcome findings, and to employ designs that account for as many threats to validity as is possible. But much of what has been written about evaluation use is simply descriptive or prescriptive. While few would recommend that we abandon common sense or experience as guides for practice, it would help to have a comprehensive model of evaluation use developed from them. To be most useful, such a model should have several characteristics.

1) It should be open-minded in defining use, and should include nonuse and misuse as well as varying levels and kinds of "good use." It should also include the possibility of explaining unintentional use, misuse, and nonuse. One possible start might be to simply

consider evaluation use as "what the user does with an evaluation."

2) Evaluation use should be conceptualized as a process or relationship rather than an object. If one were to try diagramming evaluation use as a black box, "use" would not be the product coming from the box but the undetermined process occuring within it.

3) The model should be a "real" model, one that represents the relevant components and their interrelationships with the purpose of explaining how things work. Simple rectilinear charts crossing Factor A with Factor B are not very helpful.

4) The model should treat manipulable and nonmanipulable variables differently. Factors that cannot be controlled at all, those that can be controlled only indirectly, and those that are susceptible to direct influence should be shown as such in a working model.

5) The interaction of factors should be central. In real life the political climate influences dissemination procedures, the organizational setting determines in part study design and user collaboration, and so forth. A model should explain how these interactions work.

6) The model should be parsimonious, reducing variables to the smallest number possible rather than recounting lists of empirically derived conditions. Models are useful when they simplify reality with a minimal loss of accuracy.

These are a few characteristics that would contribute to the utility of a model of evaluation use. No one would argue that evaluations are not used, and we know a great deal about what to do in order to ensure that they are. At this point there is plenty of empirical data on evaluation use and it seems appropriate to begin a little theory building.

References Cited

Alkin, M.C., Daillak, R., & White, P. (1979). Using evaluations: Does evaluation make a difference? Beverly Hills, CA: Sage.

Braskamp, A.L. & Brown, R.D. (1980). Utilization of evaluation information. San Francisco, CA: Jossey-Bass.

Brown, R.D., Newman, D.L., & Rivers, L.S. (1984). A decision-making context model for enhancing evaluation utilization. Educational Evaluation and Policy Analysis, 6, 393-400.

Burry, J. (1985). Documenting evaluation use: Guided evaluation decisionmaking. Los Angeles, CA: Center for the Study of Evaluation.(ERIC Document Reproduction Service No.ED266-154)

Carver, F.D. (1985). Ballroom dancing is not for singles. Paper presented at the meeting of the Iowa Research and Eavluation Association, Waterloo, Iowa. (ERIC Document Reproduction Service No. ED274-717)

Ciarlo, J.A. (Ed.) (1981). Utilizing evaluation: Concepts and measurement techniques. Beverly Hills, CA: Sage.

Cousins, J.B. & Leithwood, K.A. (1986). Current empirical research on evaluation utilization. Review of Educational Research, 56, 331-364.

Guba, E.G. & Lincoln, Y. S. (1981). Effective evaluation. San Francisco, CA: Jossey-Bass.

Haenn, J.F. (1980). Reasons why evaluation and testing don't inform. Paper presented at the meeting of the American Educational Research Association, Boston.

Hall, G.E. (1981). Viewing evaluation utilization as an innovation. Austin, TX: Research and Development Center for Teacher Education, The University of Texas.

Holley, F.M. (1979) Catch a falling star: Promoting the utilization of research and evaluation findings. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

King, J.A. & Pechman, E.M. (1984). Pinning a wave to the shore: Conceptualizing evaluation use in school systems. Educational Evaluation and Policy Analysis. 6, 241-251.

King, J.A. & Thompson, B. (1981) Evaluation utilization: An annotated bibliography. (ERIC Document Reproduction Service No. ED204-363)

Kleinfeld, J. & McDiarmid, G.W. (1986). Living to tell the tale: Researching politically controversial topics and communicating the findings. Educational Evaluation and Policy Analysis, 8, 393-401.

Leviton, L.A. & Hughes, E.F.X. (1981). Research in utilization of evaluations: A review and synthesis. Evaluation Review, 5, 525-548.

Meltsner, A.J. (1976). Policy analysis in the beaurocracy. Berkeley,CA: University of California Press.

Merwin, J.C. & Wiener, P.H. (1985). Evaluation: A profession? Educational Evaluation and Policy Analysis, 7, 253-259.

Patton, M.Q. (1980). Qualitative evaluation methods. Beverly Hills, CA: Sage.

Patton, M.Q. (1985). Six honest serving men for evaluation. Paper presented at the meeting of the American Educational Research Association, Chicago, IL.

Pechman, E.M. & King, J.A. (1986). Analyzing the school evaluation use process to make evaluation worth the effort. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Ruskus, J.A. & Alkin, M.C. (1984). Factors common to high-utilization evaluations. CSE Report No. 240. Los Angeles, CA: Center for the Study of Evaluation. (ERIC Document Reproduction Service No. ED251-516)

Rutman, L. (1980). Planning useful evaluations. Beverly Hills, CA: Sage.

Shapiro, J. Z. (1985). Where we are and where we need to go. Educational Evaluation and Policy Analysis, 7, 245-248.

Siegel, K. & Tuckel, P. (1985). The utilization of evaluation research: A case analysis. Evaluation Review, 9, 307-328.

Weiner, S.S., Rubin, D., & Sachse, T. (1981). Pathology in institutional structures for evaluation and a possible cure. Stanford, CA: Stanford Evaluation Consortium.

Weiss, C.H. (1981). Measuring the use of evaluation. In, J. A. Ciarlo (Ed.), Utilizing evaluation: Concepts and measurement techniques, 17-33. Beverly Hills, CA: Sage.

Wise, R.I. (1978). What we know about the decision-maker and decision settings. Paper presented at the meeting of the American Educational Research Association, Toronto, Canada.

LMH/RESEARCH/DEF-USE